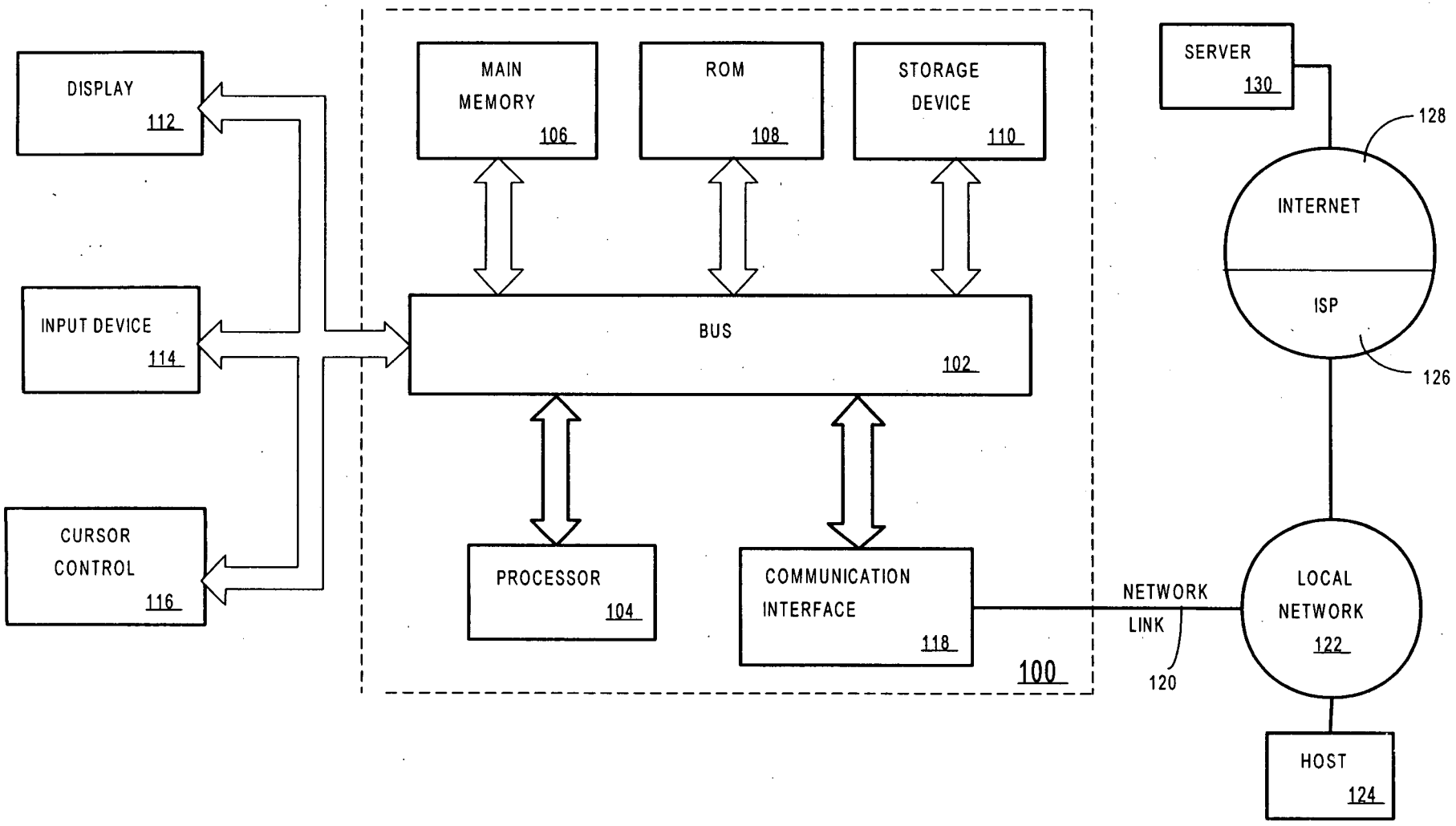
Figure 1


DISPLAY
112
INPUT DEVICE
114
CURSOR
CONTROL
116
MAIN
MEMORY
106
ROM
108
STORAGE
DEVICE
110
BUS
102
PROCESSOR
104
COMMUNICATION
INTERFACE
118
100
NETWORK
LINK
120
SERVER
130
128
INTERNET
ISP
126
LOCAL
NETWORK
122
HOST
124

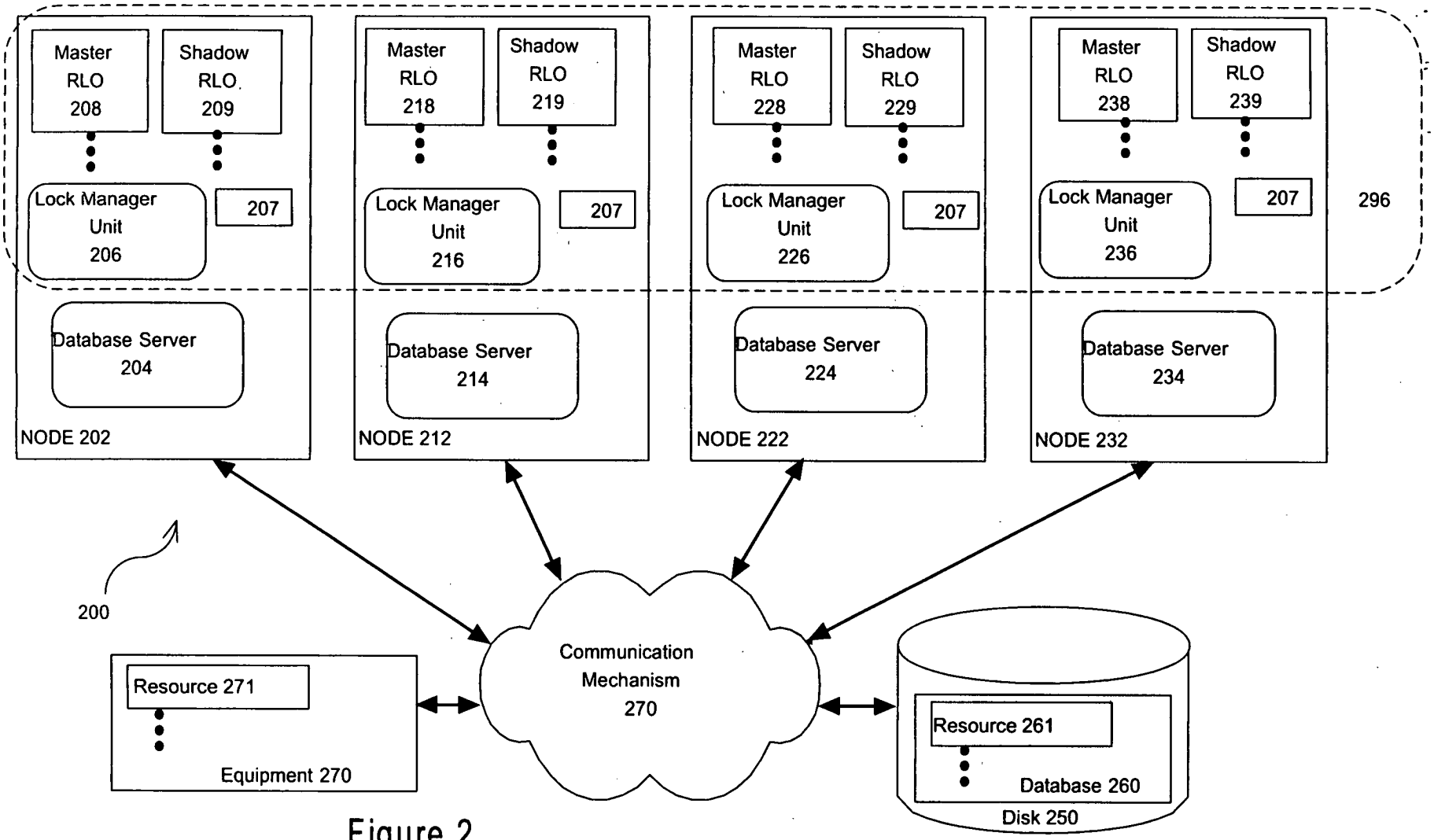

Master RLO 208
Shadow RLO 209
Lock Manager Unit 206
207
Database Server 204
NODE 202
Master RLO 218
Shadow RLO 219
Lock Manager Unit 216
207
Database Server 214
NODE 212
Master RLO 228
Shadow RLO 229
Lock Manager Unit 226
207
Database Server 224
NODE 222
Master RLO 238
Shadow RLO 239
Lock Manager Unit 236
207
Database Server 234
NODE 232
296
200
Communication Mechanism 270
Resource 271
Equipment 270
Resource 261
Database 260
Disk 250

Figure 2

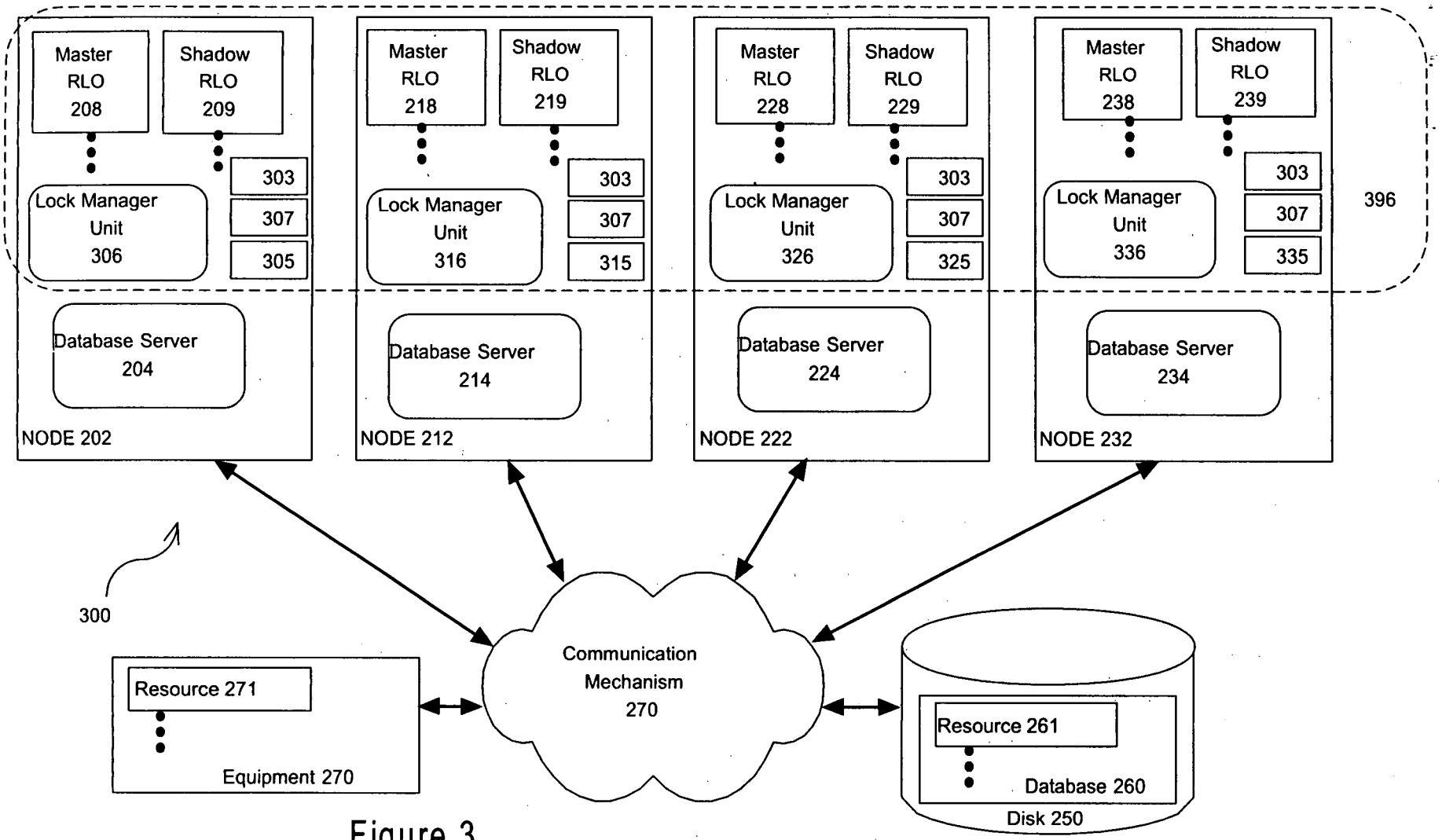

Master RLO 208
Shadow RLO 209
303
307
305
Lock Manager Unit 306
Database Server 204
NODE 202
Master RLO 218
Shadow RLO 219
303
307
315
Lock Manager Unit 316
Database Server 214
NODE 212
Master RLO 228
Shadow RLO 229
303
307
325
Lock Manager Unit 326
Database Server 224
NODE 222
Master RLO 238
Shadow RLO 239
303
307
335
Lock Manager Unit 336
Database Server 234
NODE 232
396
300
Communication Mechanism 270
Resource 271
Equipment 270
Resource 261
Database 260
Disk 250

Figure 3

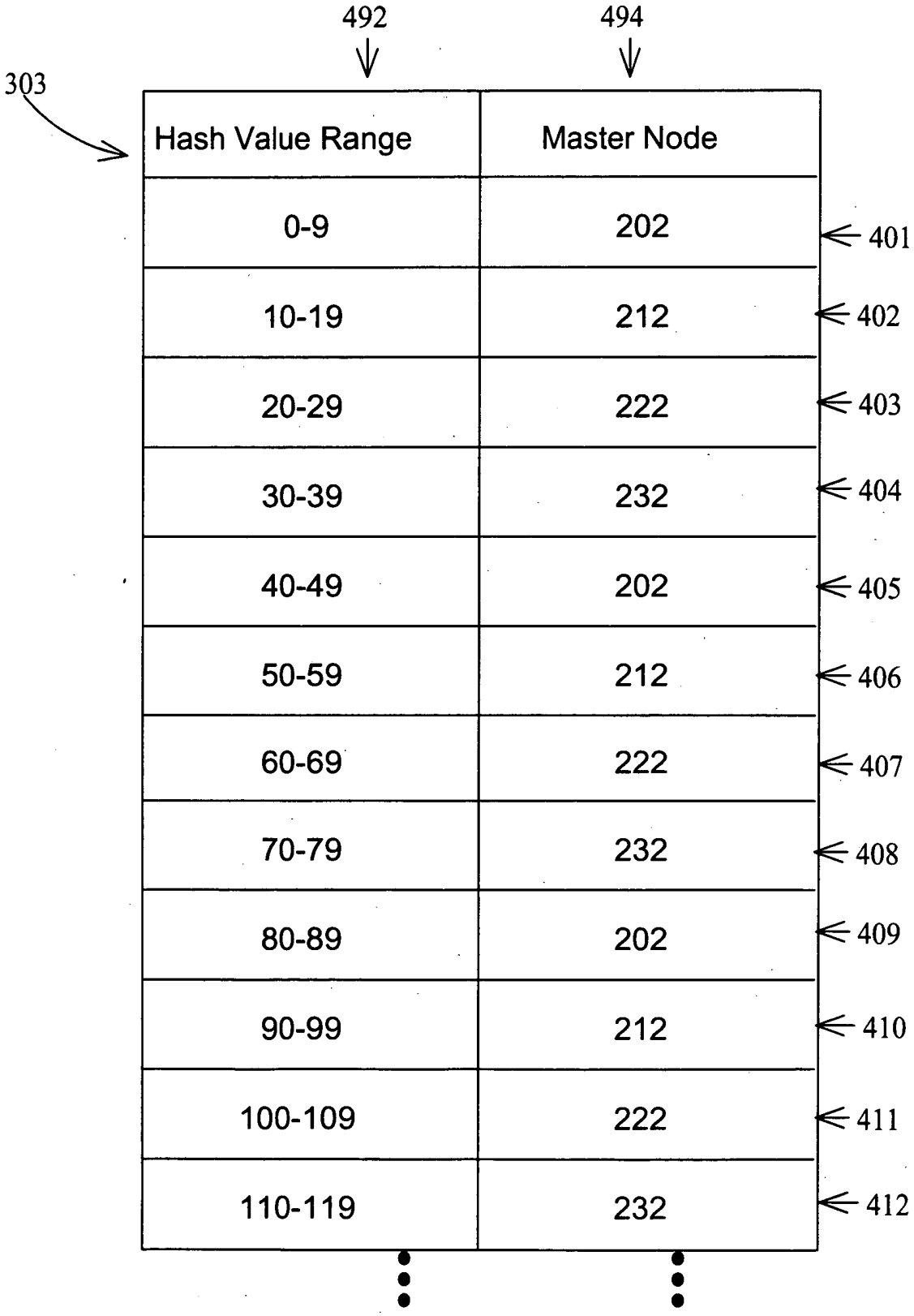

492
494
303
Hash Value Range
Master Node
0-9
202
401
10-19
212
402
20-29
222
403
30-39
232
404
40-49
202
405
50-59
212
406
60-69
222
407
70-79
232
408
80-89
202
409
90-99
212
410
100-109
222
411
110-119
232
412

Figure 4A

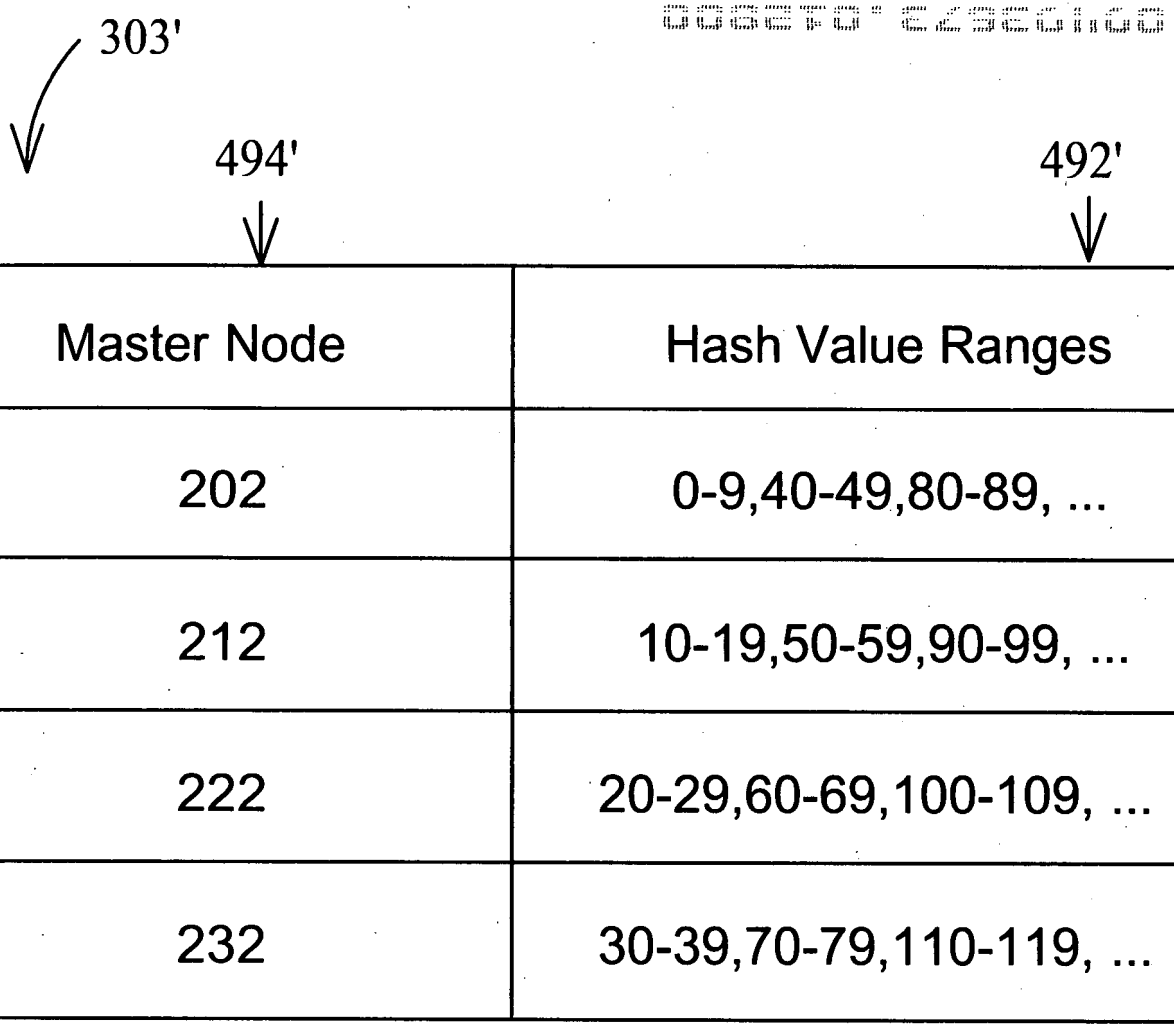

303'
494'
492'
Master Node
Hash Value Ranges
202
0-9,40-49,80-89, ...
421
212
10-19,50-59,90-99, ...
422
222
20-29,60-69,100-109, ...
423
232
30-39,70-79,110-119, ...
424

Figure 4B

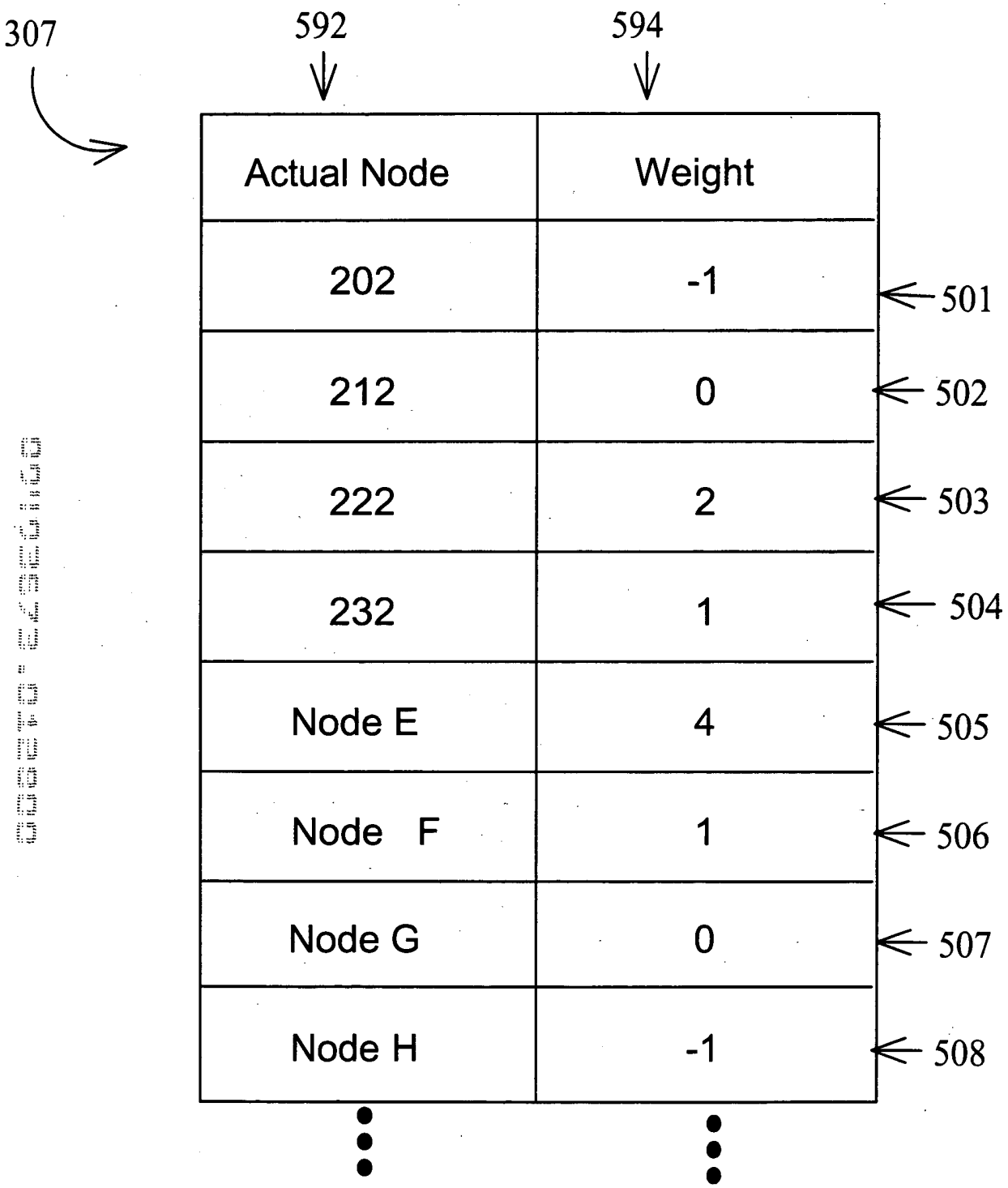

307
592
594
Actual Node
Weight
202
-1
501
212
0
502
222
2
503
232
1
504
Node E
4
505
Node F
1
506
Node G
0
507
Node H
-1
508

Figure 5

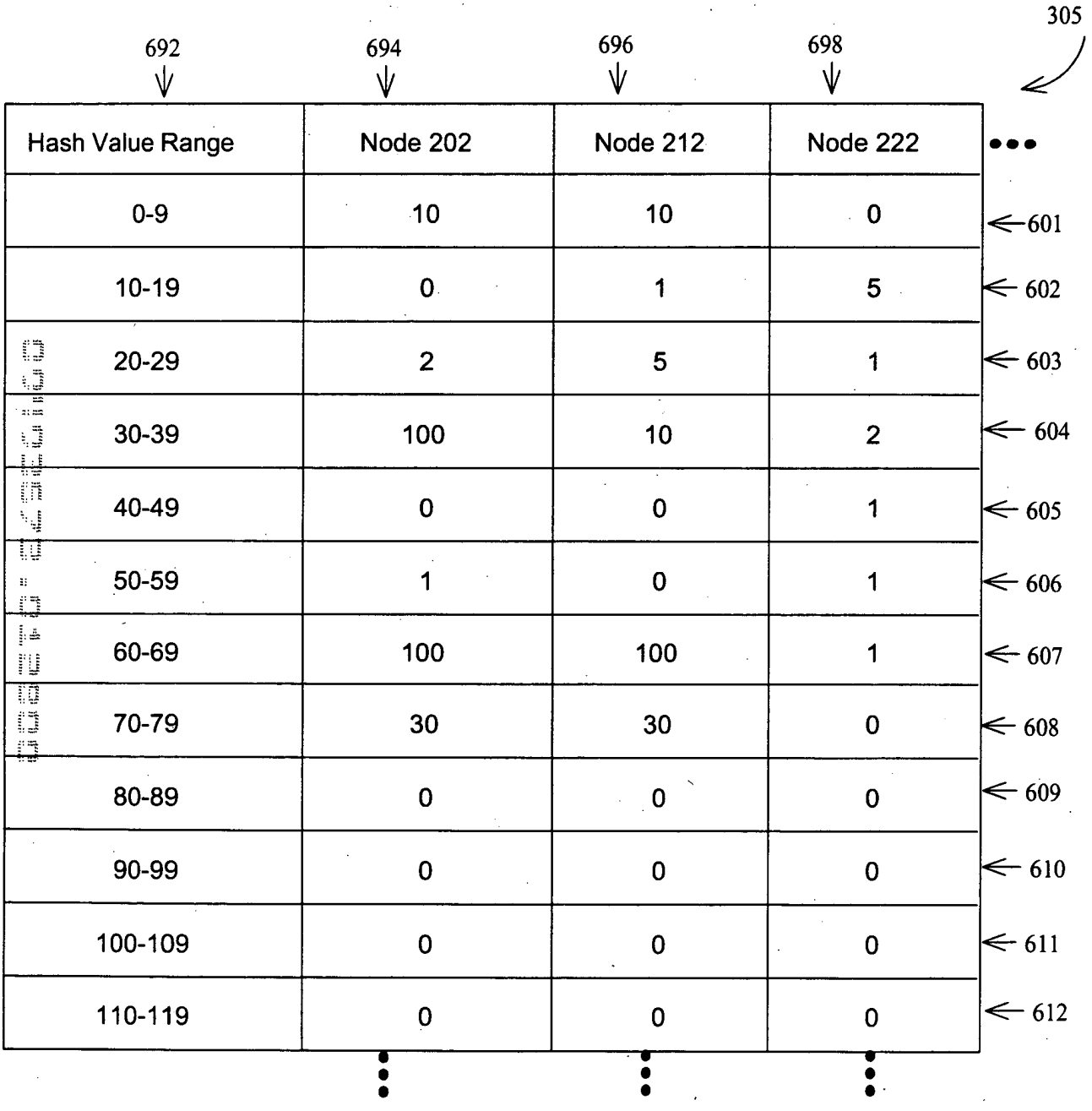

305
692
694
696
698
Hash Value Range
Node 202
Node 212
Node 222
0-9
10
10
0
601
10-19
0
1
5
602
20-29
2
5
1
603
30-39
100
10
2
604
40-49
0
0
1
605
50-59
1
0
1
606
60-69
100
100
1
607
70-79
30
30
0
608
80-89
0
0
0
609
90-99
0
0
0
610
100-109
0
0
0
611
110-119
0
0
0
612

Figure 6

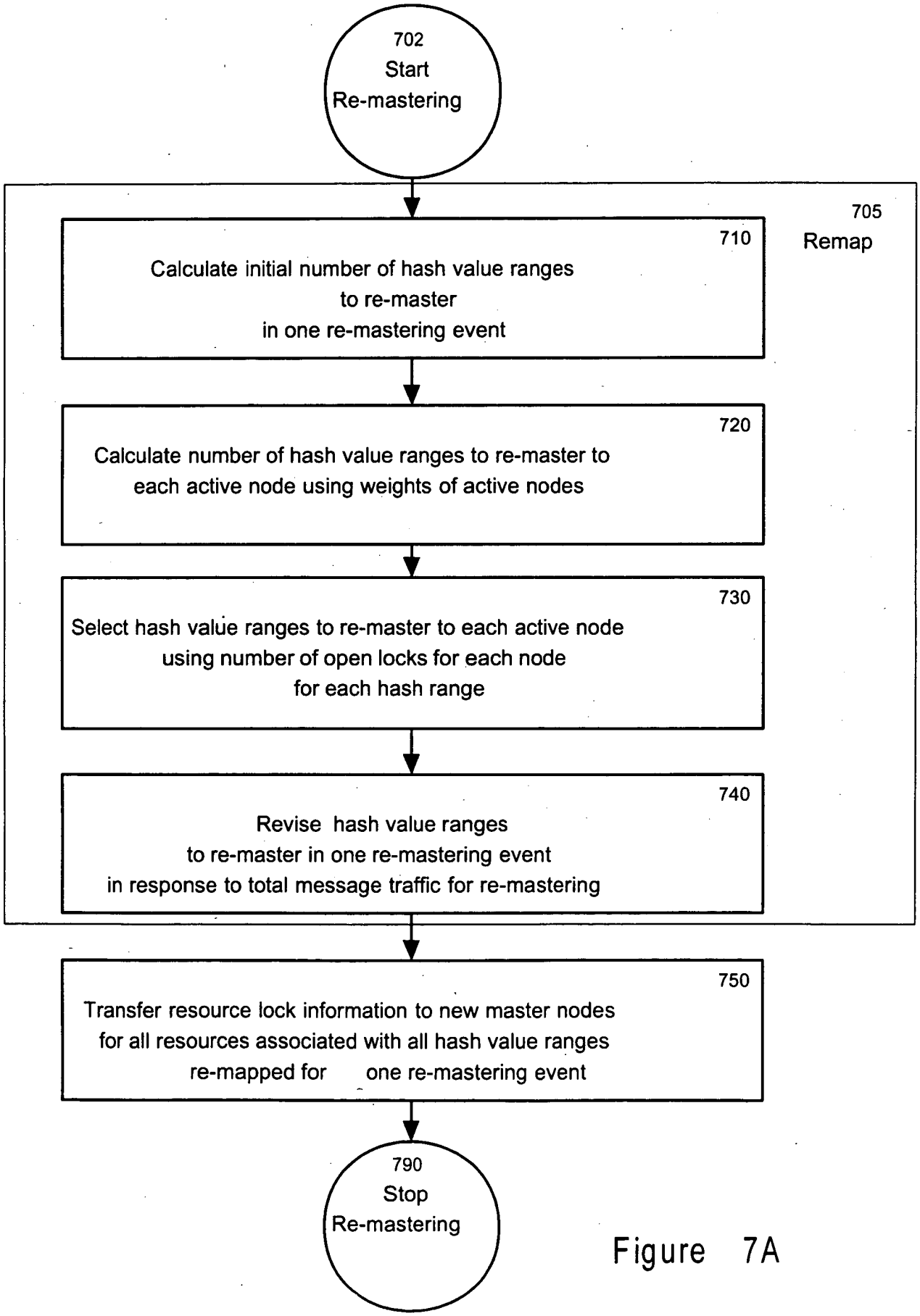

702
Start
Re-mastering
705
Remap
710
Calculate initial number of hash value ranges
to re-master
in one re-mastering event
720
Calculate number of hash value ranges to re-master to
each active node using weights of active nodes
730
Select hash value ranges to re-master to each active node
using number of open locks for each node
for each hash range
740
Revise hash value ranges
to re-master in one re-mastering event
in response to total message traffic for re-mastering
750
Transfer resource lock information to new master nodes
for all resources associated with all hash value ranges
re-mapped for one re-mastering event
790
Stop
Re-mastering

Figure 7A

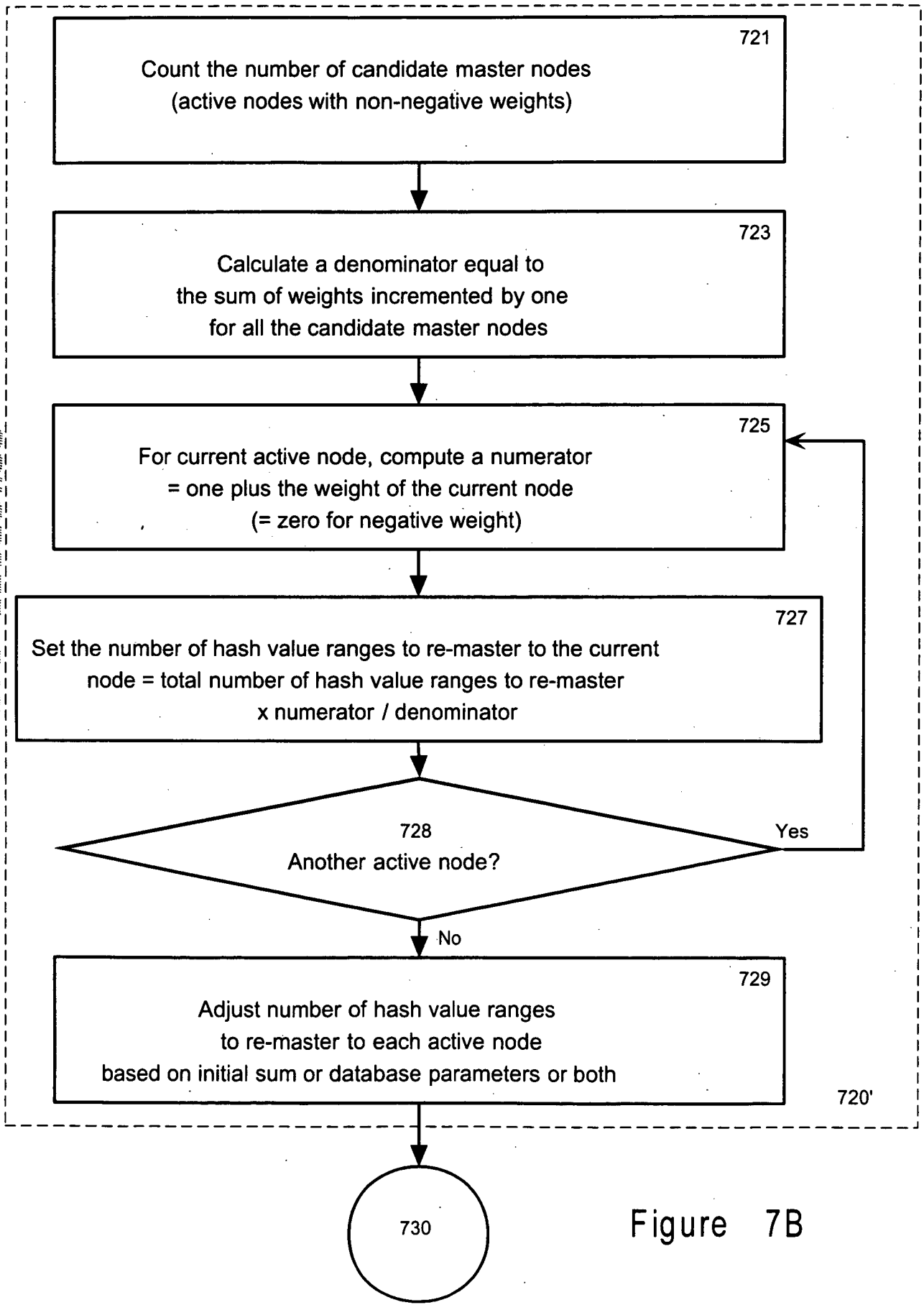

721
Count the number of candidate master nodes
(active nodes with non-negative weights)
723
Calculate a denominator equal to
the sum of weights incremented by one
for all the candidate master nodes
725
For current active node, compute a numerator
= one plus the weight of the current node
(= zero for negative weight)
727
Set the number of hash value ranges to re-master to the current
node = total number of hash value ranges to re-master
x numerator / denominator
728
Another active node?
Yes
No
729
Adjust number of hash value ranges
to re-master to each active node
based on initial sum or database parameters or both
720'
730

Figure 7B

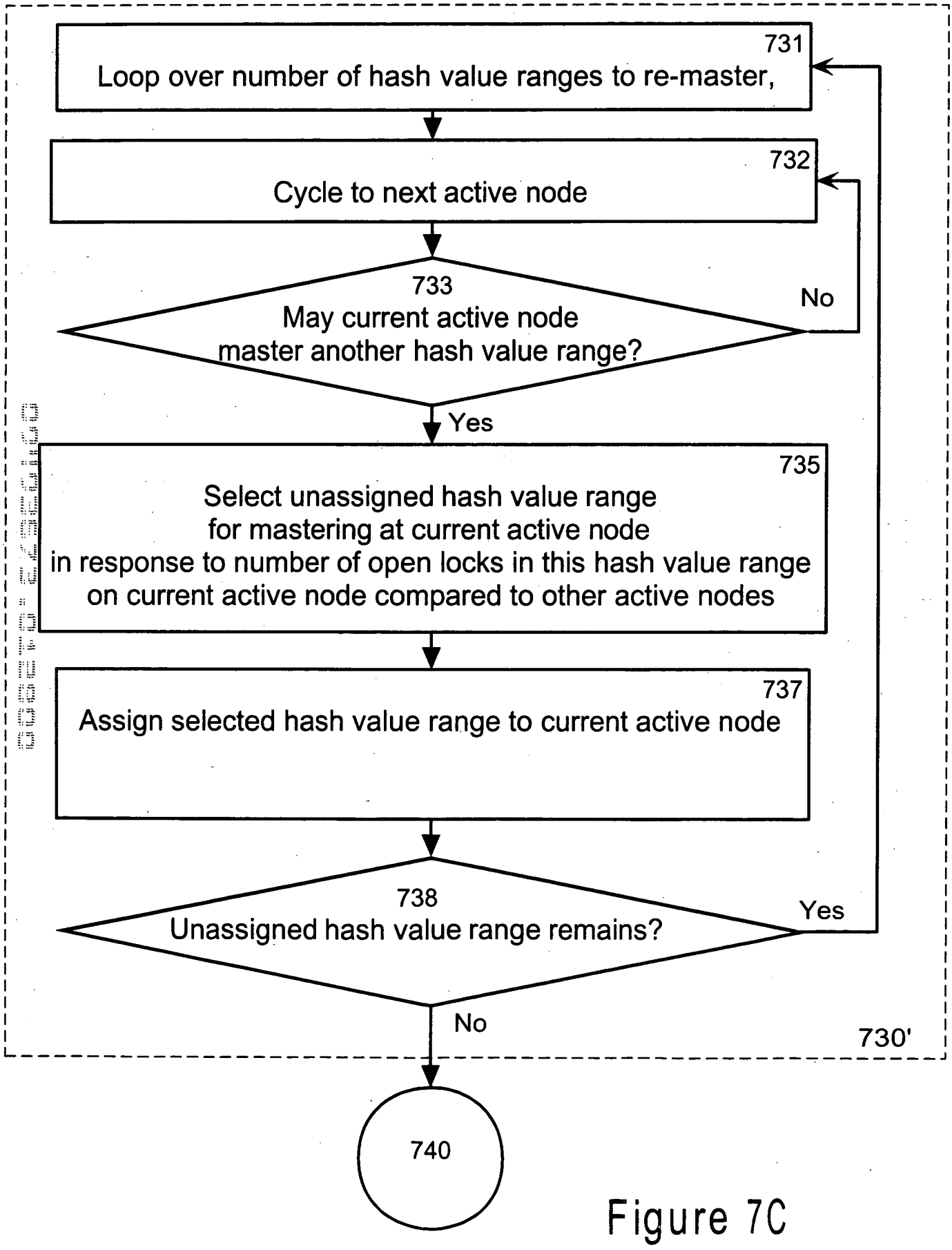

731
Loop over number of hash value ranges to re-master,
732
Cycle to next active node
733
May current active node master another hash value range?
No
Yes
735
Select unassigned hash value range for mastering at current active node in response to number of open locks in this hash value range on current active node compared to other active nodes
737
Assign selected hash value range to current active node
738
Unassigned hash value range remains?
Yes
No
730'
740

Figure 7C

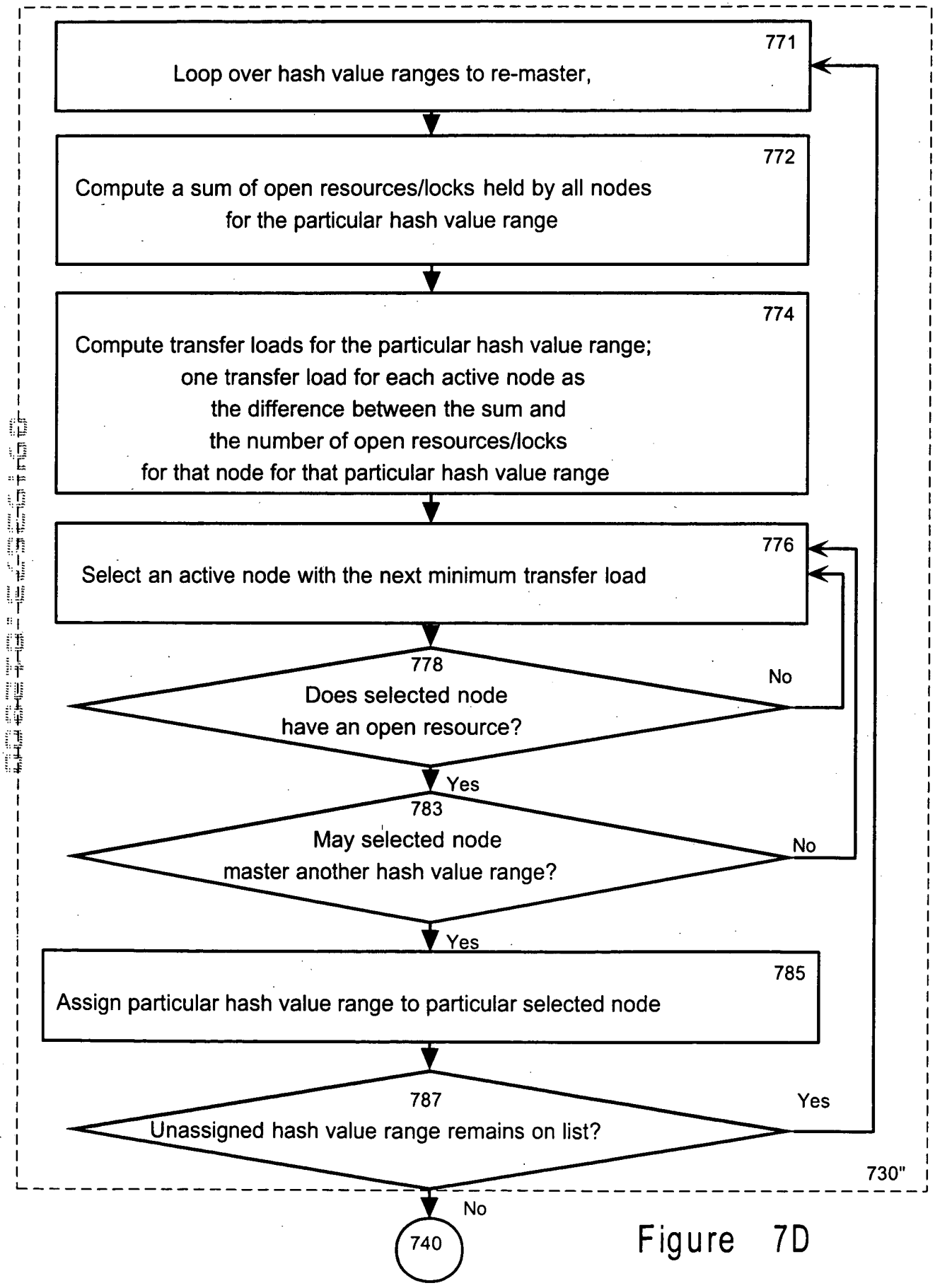

771
Loop over hash value ranges to re-master,
772
Compute a sum of open resources/locks held by all nodes for the particular hash value range
774
Compute transfer loads for the particular hash value range; one transfer load for each active node as the difference between the sum and the number of open resources/locks for that node for that particular hash value range
776
Select an active node with the next minimum transfer load
778
Does selected node have an open resource?
No
Yes
783
May selected node master another hash value range?
No
Yes
785
Assign particular hash value range to particular selected node
787
Unassigned hash value range remains on list?
Yes
730"
No
740

Figure 7D

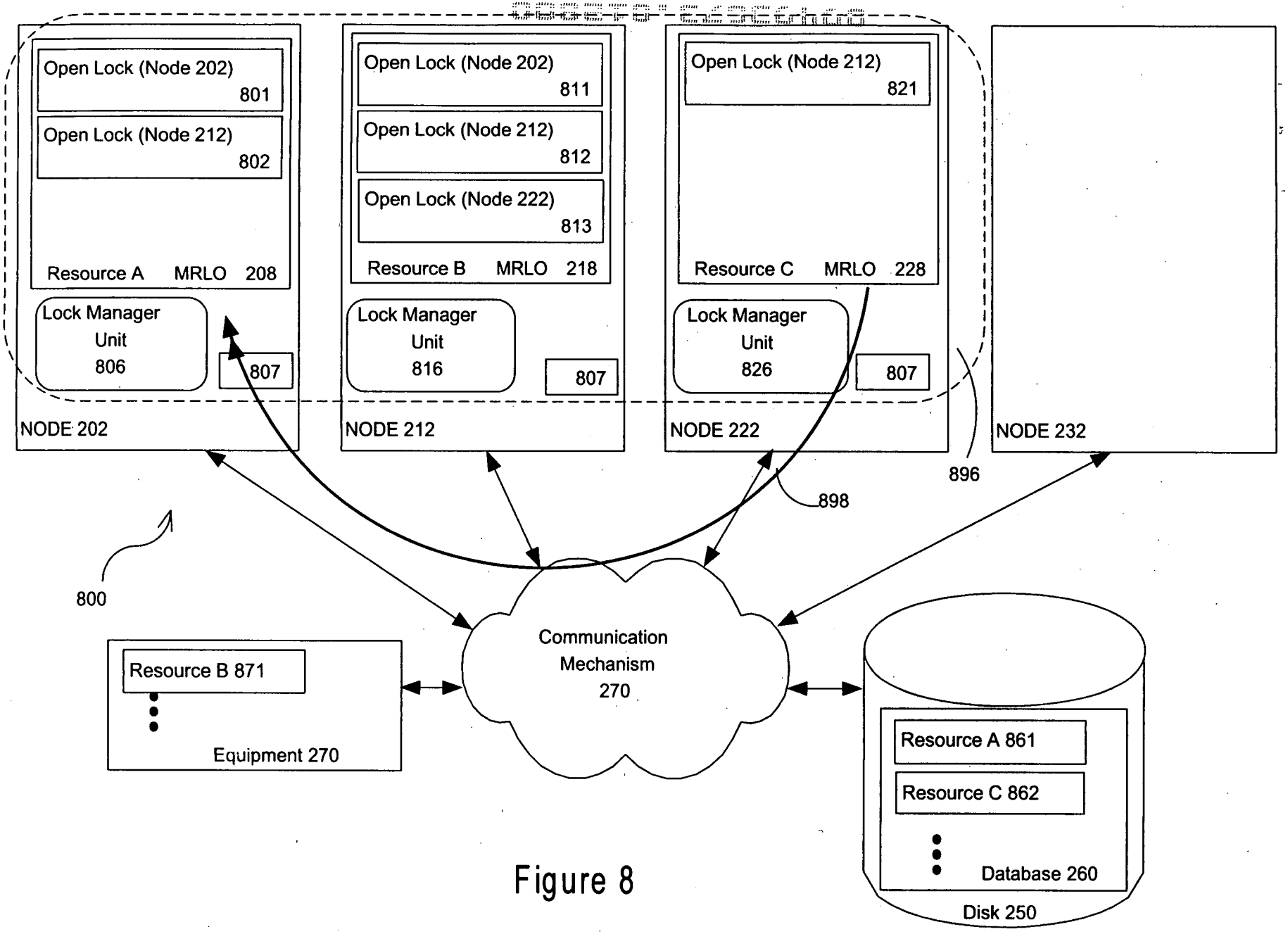

Open Lock (Node 202)
801
Open Lock (Node 212)
802
Resource A MRLO 208
Lock Manager Unit 806
807
NODE 202
Open Lock (Node 202)
811
Open Lock (Node 212)
812
Open Lock (Node 222)
813
Resource B MRLO 218
Lock Manager Unit 816
807
NODE 212
Open Lock (Node 212)
821
Resource C MRLO 228
Lock Manager Unit 826
807
NODE 222
NODE 232
896
898
800
Communication Mechanism 270
Resource B 871
Equipment 270
Resource A 861
Resource C 862
Database 260
Disk 250

Figure 8

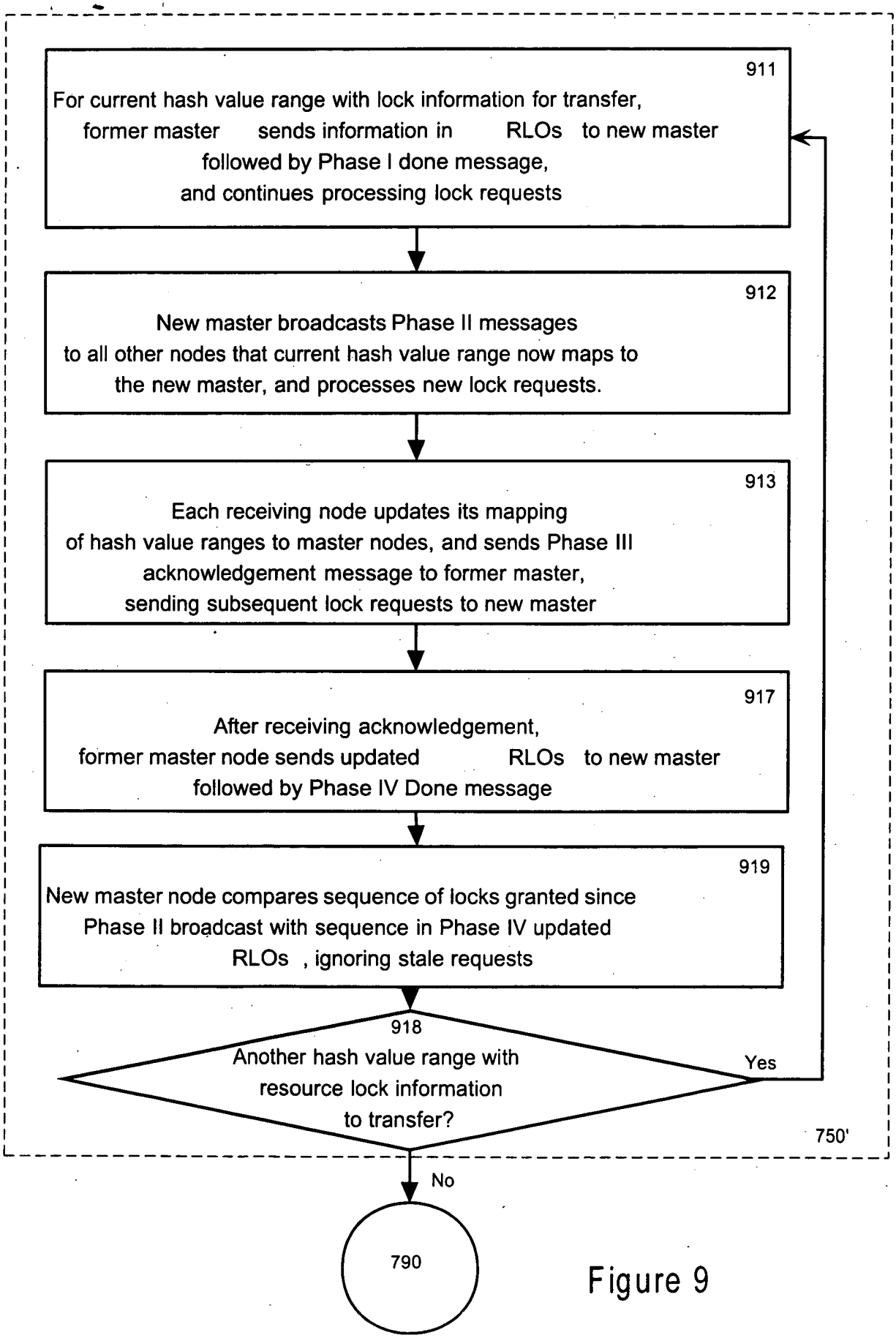

911
For current hash value range with lock information for transfer, former master sends information in RLOs to new master followed by Phase I done message, and continues processing lock requests
912
New master broadcasts Phase II messages to all other nodes that current hash value range now maps to the new master, and processes new lock requests.
913
Each receiving node updates its mapping of hash value ranges to master nodes, and sends Phase III acknowledgement message to former master, sending subsequent lock requests to new master
917
After receiving acknowledgement, former master node sends updated RLOs to new master followed by Phase IV Done message
919
New master node compares sequence of locks granted since Phase II broadcast with sequence in Phase IV updated RLOs , ignoring stale requests
918
Another hash value range with resource lock information to transfer?
Yes
No
750'
790

Figure 9

Figure 10A

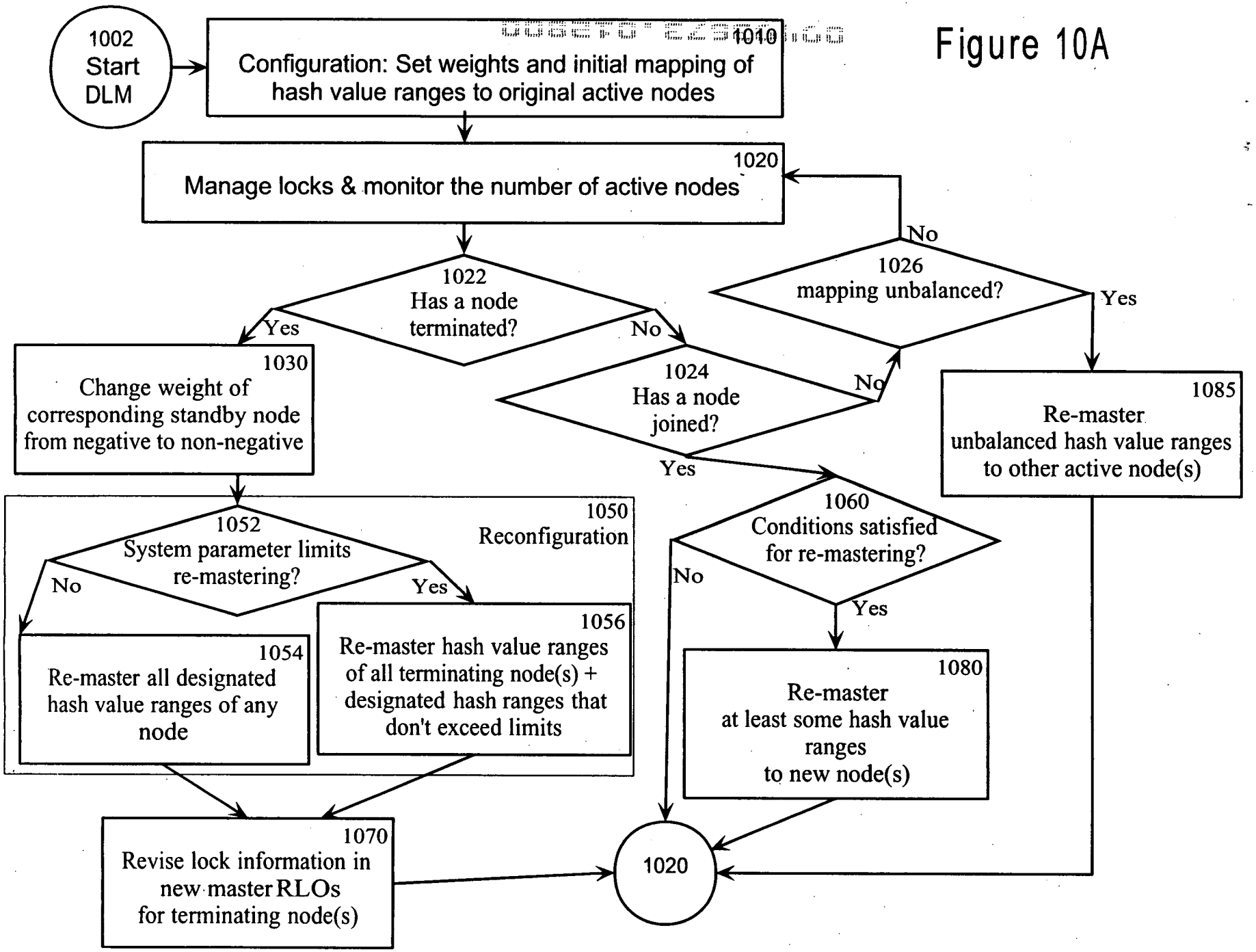
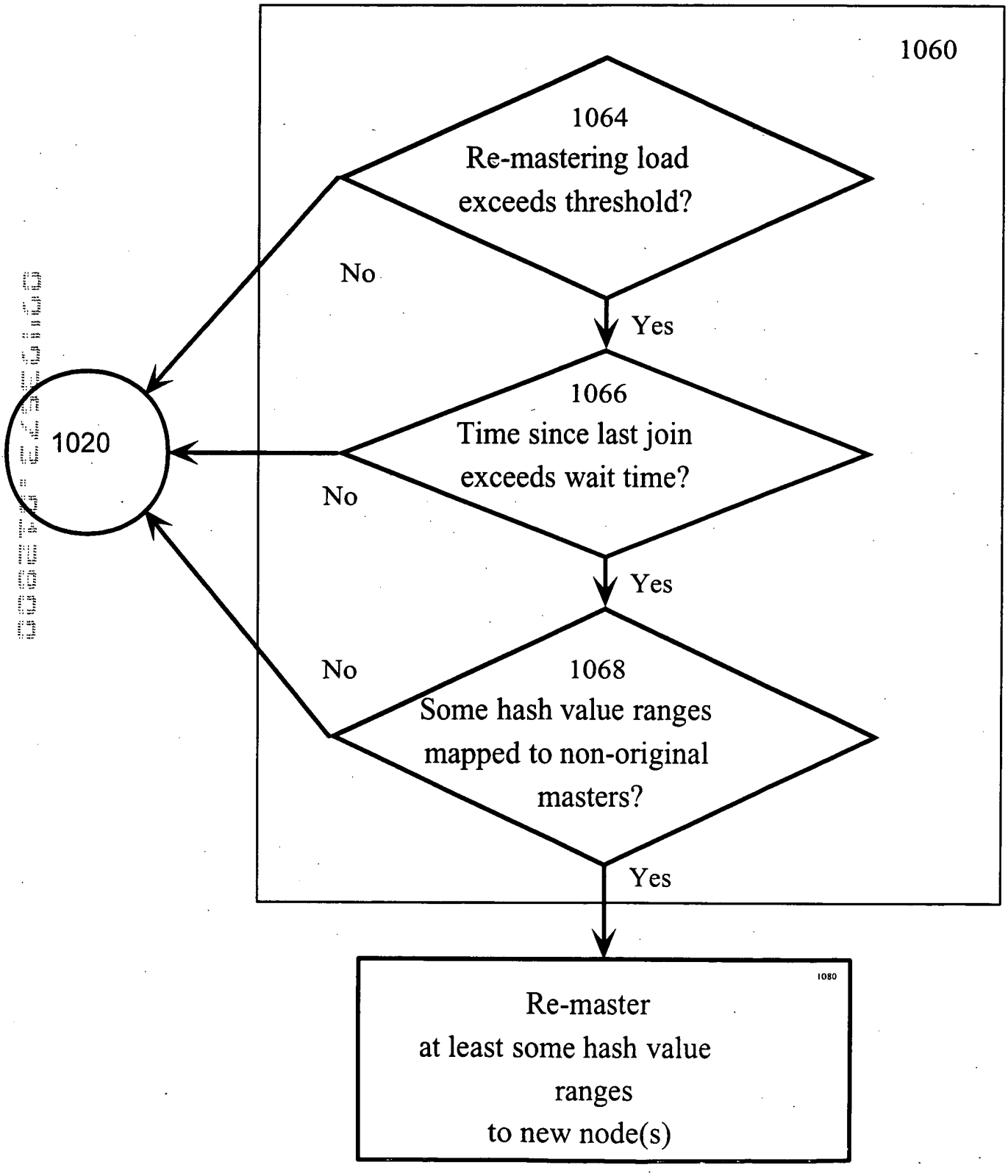

1002
Start
DLM
1010
Configuration: Set weights and initial mapping of hash value ranges to original active nodes
1020
Manage locks & monitor the number of active nodes
1022
Has a node terminated?
Yes
No
1024
Has a node joined?
No
Yes
1026
mapping unbalanced?
No
Yes
1030
Change weight of corresponding standby node from negative to non-negative
1085
Re-master unbalanced hash value ranges to other active node(s)
1050
Reconfiguration
1052
System parameter limits re-mastering?
No
Yes
1054
Re-master all designated hash value ranges of any node
1056
Re-master hash value ranges of all terminating node(s) + designated hash ranges that don't exceed limits
1060
Conditions satisfied for re-mastering?
No
Yes
1080
Re-master at least some hash value ranges to new node(s)
1070
Revise lock information in new master RLOs for terminating node(s)
1020

Figure 10B


1060
1064
Re-mastering load exceeds threshold?
No
Yes
1066
Time since last join exceeds wait time?
No
Yes
1020
1068
Some hash value ranges mapped to non-original masters?
No
Yes
1080
Re-master at least some hash value ranges to new node(s)

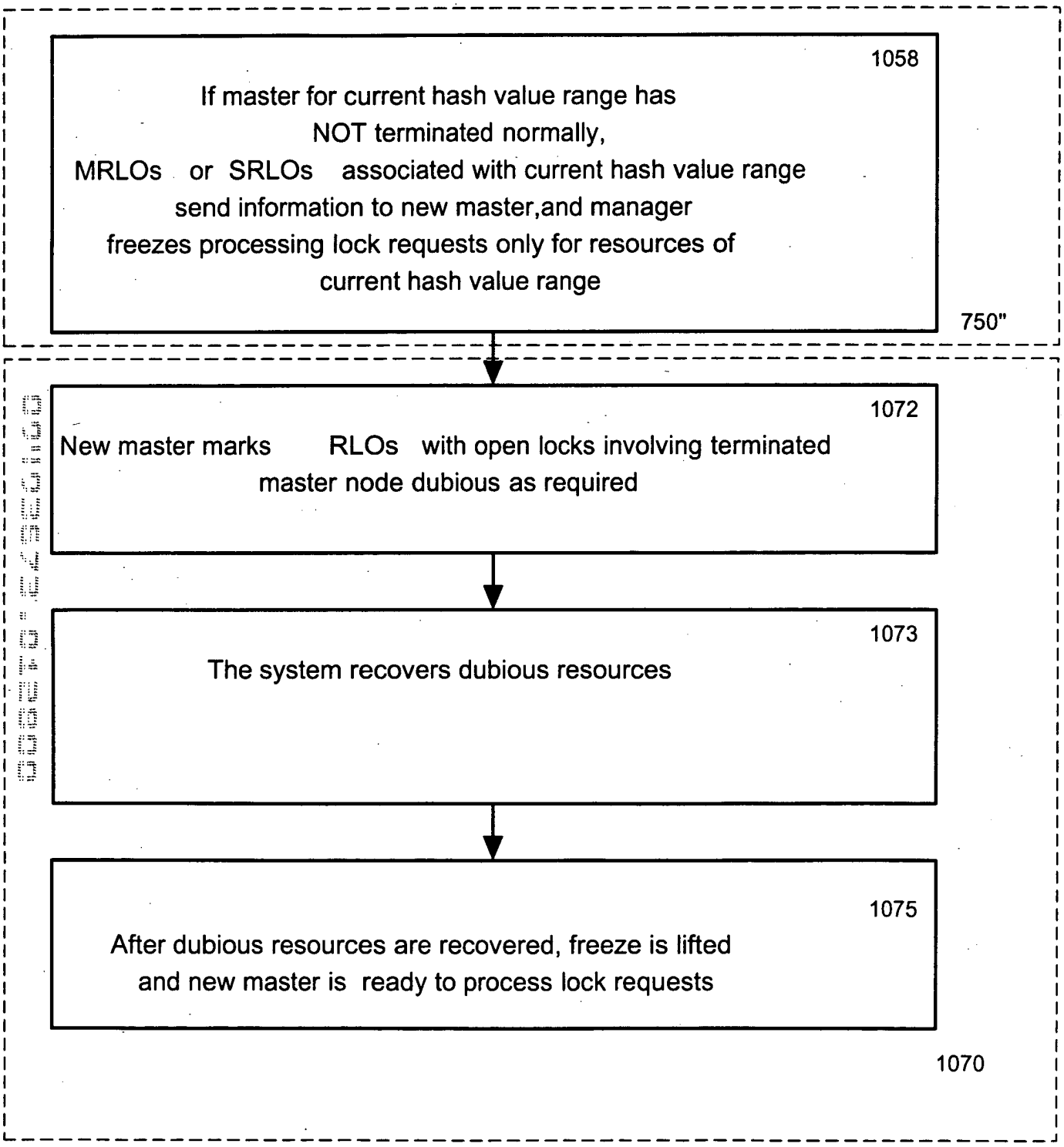

1058
If master for current hash value range has NOT terminated normally, MRLOs or SRLOs associated with current hash value range send information to new master,and manager freezes processing lock requests only for resources of current hash value range
750"
1072
New master marks RLOs with open locks involving terminated master node dubious as required
1073
The system recovers dubious resources
1075
After dubious resources are recovered, freeze is lifted and new master is ready to process lock requests
1070

Figure 10C